Journal of Econometrics 238 (2024) 105546

Contents lists available at ScienceDirect

Journal of Econometrics

journal homepage: www.elsevier.com/locate/jeconom

Simultaneously Incomplete and Incoherent (SII) Dynamic LDV Models: With an Application to Financing Constraints and Firms' Decision to Innovate

Vassilis Hajivassiliou [*,1,a], Frédérique Savignac [1,b]

[a] *Dep of Econ and FMG, LSE, London WC2A 2AE, England Department of Economics, LSE and FMG United Kingdom*
[b] *Banque de France France*

ARTICLE INFO

JEL classification:
C51
C52
C15

Keywords:
Financing Constraints
Innovation
Dynamic Limited Dependent Variable Models
Joint Bivariate Probit Model
Econometric Coherency and Completeness Conditions
State Dependence

ABSTRACT

We develop novel methods for establishing coherency and completeness conditions in Static and Dynamic Limited Dependent Variables (LDV) Models. We characterize the two distinct problems as "empty-region" incoherency and "overlap-region" incoherency or incompleteness and show that the two properties can co-exist. We focus on the class of models that can be Simultaneously Incomplete and Incoherent (SII). We propose estimation strategies based on Conditional Maximum Likelihood Estimation (CMLE) for simultaneous dynamic LDV models without imposing recursivity. Point identification is achieved through sign-restrictions on parameters or other prior assumptions that complete the underlying data process. Using as modelling framework the Panel Bivariate Probit model with State Dependence, we analyse the impact of financing constraints on innovation: ceteris paribus, a firm facing binding finance constraints is substantially less likely to undertake innovation, while the probability that a firm encounters a binding finance constraint more than doubles if the firm is innovative. In addition, a strong role for state dependence in dynamic versions of our models is established.

1. Introduction

In this paper we investigate the fundamental identification issues of *coherency* and *completeness* of Limited Dependent Variable (LDV) models with endogeneity and flexible temporal and contemporaneous correlations in the unobservables. We focus on the class of models that can be Simultaneously Incomplete and Incoherent (SII), which was largely ignored by the existing literature.

Our main contributions can be summarized as follows: in Section 2, we define and distinguish two types of identification problems

* Corresponding author.
E-mail address: vassilis@econ.lse.ac.uk (V. Hajivassiliou).

[1] We have benefitted from constructive comments by participants at seminars at the London School of Economics, Cemmap/UCL, CERGE, and at the universities of Oxford, City of London, Marseilles, Milan, Toulouse, Cambridge, Cyprus, and York. Early ideas on which this work is based were presented at the 2006 Conference on Panel Data at the University of Cambridge, the 2006 Corporate Finance Conference in London, and the IGIER 2008 Conference in Capri. We thank in particular Richard Blundell, Christian Bontemps, Andrew Chesher, Francesca Cornelli, Antoine Faure-Grimaud, Denis Gromb, Tatiana Komarova, Guy Laroque, Claire Lelarge, Thierry Magnac, Alex Michaelides, Peter Robinson, John Sutton, Elie Tamer, Alain Trognon, Dimitri Vayanos, Paolo Volpin, and Frank Windjmeier. We also thank Ryan Giordano for expert research assistance. We are indebted to the Co-Editor and two anonymous referees for very constructive comments and suggestions. The opinions expressed in this paper are not necessarily those of the Banque de France or the Eurosystem. The usual disclaimer applies for all remaining errors.

https://doi.org/10.1016/j.jeconom.2023.105546
Received 1 March 2021; Received in revised form 27 January 2023; Accepted 30 March 2023

termed "empty-region incoherency" and "completeness or overlap-region incoherency." This distinction clears up the past literature where confusing and frequently conflicting definitions were used. In Section 3, we discuss the class of Simultaneously Incomplete and Incoherent (SII) models, which exhibit simultaneously *both* types of incoherency. We thus establish that incoherency and incompleteness are not a model's either/or properties but they could co-exist. Furthermore, while the traditional approaches focus on *sufficient* conditions for coherency, our methods prove that they *may not be necessary*. We discuss the similarities with and differences from the existing literature in Section 4.

Our third contribution, presented in Section 5, is to establish coherency and completeness of static and dynamic LDV models without imposing recursivity. We achieve this through the use of (a) endogeneity in terms of latent variables and/or (b) sign restrictions on model parameters. The issues and solutions are first presented through a Graphical approach, while in Section 6 we develop an alternative analysis which we term the DGP approach. We employ this approach to solve considerably more complicated LDV models, especially those that contain intertemporal endogeneity.

The novel methods we develop, which have intuitive interpretations and are easy to implement and generalize, lead to estimation strategies based on Conditional Maximum Likelihood Estimation (CMLE) for simultaneous LDV models. The methods also allow us to establish the coherency of several Dynamic LDV models, for which it was impossible to determine whether they were coherent and/or incomplete using traditional methods. We then establish for the first time the coherency and completeness of the Panel Univariate Probit model with State Dependence and that of the Panel Bivariate Probit model with State Dependence. Extensions and other issues are discussed in Section 7.

In Section 8, we propose an empirical application to quantify the interactions between innovation by firms and the financial constraints they may face. We use as the modelling framework the Panel Bivariate Probit model with State Dependence.

Our empirical results are quite striking: ceteris paribus, we estimate that a firm that faces a binding finance constraint is approximately 30% less likely to undertake innovation, while the probability that a firm encounters a binding finance constraint more than doubles if the firm is classified as innovative. Finally, we establish a strong role for state dependence: firms tend to innovate continuously rather than occasionally and past financial difficulties are correlated with the present ones even after conditioning on important firm characteristics. We thus establish the value of our novel approach, which for the first time eliminates the need to assume model recursivity Section 9 concludes.

2. Coherency and completeness as distinct, one-at-a-time problems

We focus our analysis by using the Simultaneous LDV Model with Two Binary Responses. In this model, limited dependent variables y_1 and y_2 are jointly determined through filter functions $\tau_1(\cdot)$ and $\tau_2(\cdot)$ operating on latent variables y_1^* and y_2^* respectively:

$$y_{1it} = \tau_1\left(y_{1it}^* \equiv \left[h_1\left(x_{1it}'\beta_1, y_{2it}\gamma\right) + \epsilon_{1it}\right]\right) \tag{1}$$

$$y_{2it} = \tau_2\left(y_{2it}^* \equiv \left[h_2\left(x_{2it}'\beta_2, y_{1it}\delta\right) + \epsilon_{2it}\right]\right) \tag{2}$$

The (possibly non-linear) functions $h_1(\cdot)$ and $h_2(\cdot)$ are known up to parameter vectors β_1 and β_2 and the two interaction coefficients γ and δ. The interaction terms $y_{2it}\gamma$ and $y_{1it}\delta$ appear in the respective latent variables y_{1it}^* and y_{2it}^*. Let x_{1it} and x_{2it} denote the vectors of exogenous factors for each side. The parameter vector to be estimated is $\theta \equiv (\beta_1', \beta_2', \gamma, \delta, \sigma_1^2, \sigma_2^2, \rho)'$ where $\rho \equiv \mathit{correlation}(\epsilon_{1it}, \epsilon_{2it})$. In the most general case, the sample is a panel data set indexed by $i = 1, \cdots, N$ and $t = 1, \cdots, T_i$.

The existing econometric literature has established as the typical coherency condition to be: $\gamma \cdot \delta = 0$, i.e., no reverse interaction terms are allowed among the two endogenous variables. This condition, which is termed "recursivity," is sufficient for the joint distribution $(y_{1it}, y_{2it}|x_1, x_2, \theta)$ to be well-specified.

The first case we focus on is the binary threshold crossing response model where: $\tau_j(z) \equiv \mathbf{1}(z > 0)$ where $\mathbf{1}(z > 0)$ is the *indicator function* defined by:

$$\mathbf{1}(z > 0) \equiv \begin{cases} 1 & \text{if } z > 0 \\ 0 & \text{if } z \le 0 \end{cases}.$$

In terms of the two latent variables y_1^* and y_2^* and the observed binary indicators y_1 and y_2, and suppressing the observation indices, we have:

$$y_1 = \begin{cases} 1 & \mathit{if} \quad y_1^* \equiv x_1'\beta_1 + \gamma y_2 + \epsilon_1 > 0 \\ 0 & \mathit{if} \quad y_1^* \equiv x_1'\beta_1 + \gamma y_2 + \epsilon_1 \le 0 \end{cases} \tag{3}$$

$$y_2 = \begin{cases} 1 & \mathit{if} \quad y_2^* \equiv x_2'\beta_2 + \delta y_1 + \epsilon_2 > 0 \\ 0 & \mathit{if} \quad y_2^* \equiv x_2'\beta_2 + \delta y_1 + \epsilon_2 \le 0 \end{cases} \tag{4}$$

In the empirical application of Section 8.2, we employed this model to study the impact of financing constraints on a firm's decision

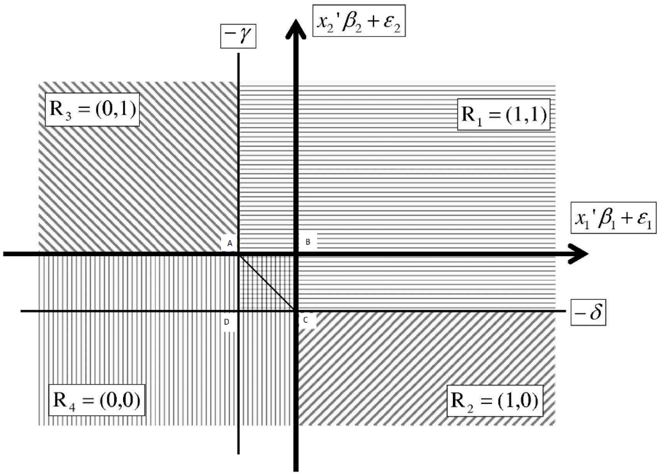

Fig. 1. Case 1: $\gamma > 0, \delta > 0$ Innovation and Finance Constraint bivariate binary probit in latent variables space. Four implied regimes. Region of Incoherency is of the overlap type (cross-hatched).

and ability to innovate in panel data.[2]

The specific version of joint binary probit model becomes:[3]

$$I_{it} = \begin{cases} 1 & if \quad I_{it}^* \equiv x_{it}^I\beta^I + \gamma F_{it} + \epsilon_{it}^I > 0 \\ 0 & if \quad I_{it}^* \equiv x_{it}^I\beta^I + \gamma F_{it} + \epsilon_{it}^I \leq 0 \end{cases} \tag{5}$$

$$F_{it} = \begin{cases} 1 & if \quad F_{it}^* \equiv x_{it}^F\beta^F + \delta I_{it} + \epsilon_{it}^F > 0 \\ 0 & if \quad F_{it}^* \equiv x_{it}^F\beta^F + \delta I_{it} + \epsilon_{it}^F \leq 0 \end{cases} \tag{6}$$

Analytically, $(I_{it}, F_{it}) \in \{(1,1),(1,0),(0,1),(0,0)\}$ such that:

(I_{it}, F_{it})	I_{it}^*		F_{it}^*
(1,1)	$x_{1it}'\beta_1 + \gamma + \epsilon_{1it} > 0$	,	$x_{2it}'\beta_2 + \delta + \epsilon_{2it} > 0$
(1,0)	$x_{1it}'\beta_1 + \epsilon_{1it} > 0$	,	$x_{2it}'\beta_2 + \delta + \epsilon_{2it} < 0$
(0,1)	$x_{1it}'\beta_1 + \gamma + \epsilon_{1it} < 0$	,	$x_{2it}'\beta_2 + \epsilon_{2it} > 0$
(0,0)	$x_{1it}'\beta_1 + \epsilon_{1it} < 0$	,	$x_{2it}'\beta_2 + \epsilon_{2it} < 0$

For a typical *it* observation, the probability $Prob(I_{it}, F_{it}|X, \theta)$ is thus characterized by the constraints on the unobservables: $(a^I, a^F)' < (\epsilon^I, \epsilon^F)' < (b^I, b^F)'$ defined by:

I_{it}	F_{it}	a^I	b^I	a^F	b^F
1	1	$-x_{it}^I\beta^I - \gamma$	∞	$-x_{it}^F\beta^F - \delta$	∞
1	0	$-x_{it}^I\beta^I$	∞	$-\infty$	$-x_{it}^F\beta^F - \delta$
0	1	$-\infty$	$-x_{it}^I\beta^I - \gamma$	$-x_{it}^F\beta^F$	∞
0	0	$-\infty$	$-x_{it}^I\beta^I$	$-\infty$	$-x_{it}^F\beta^F$

In general, in the absence of coherency conditions, there will be *overlaps* and/or *gaps* in the domain of $(\epsilon_{1it} + x_{1it}'\beta_1, \epsilon_{2it} + x_{2it}'\beta_2)$. Theoretically, Triangularity is an assumption that imposes the Coherency and Completeness conditions through $\gamma = 0$ and/or $\delta = 0$

[2] A related application of this setup in International Finance is the Banking and Currency Crises Model of Falcetti and Tudela (2008) where (C_{it}, B_{it}) refer to Currency and Banking Crises respectively. Their model is recursive, in that Currency crises are allowed to depend on Banking crises but not vice-versa.

[3] For the first equation, I^* is used for the latent and I for the observed LDV as a mnemonic to the *Innovation* side of the model of Section 8.2 below. Similarly, for the second equation we use F^* and F as a mnemonic to *Financing Constraints*.

restrictions. In typical applications, these restrictions are unrealistically stringent and should be relaxed.[4]

Prior sign restrictions on the parameters, e.g., $\gamma^*\delta < 0$, would be entirely analogous in achieving identification.[5] Obviously, there exist **four cases** based on the signs of γ and δ. In our empirical analysis below, there is a behavioural/economic motivation in terms of the *signs* of the interaction terms.[6] We now analyze the theoretical implications of these four possible cases:

2.1. Cases 1 and 2: $\gamma \cdot \delta > 0$ — overlapping regions incoherency, i.e., incompleteness

When the interaction coefficients γ and δ are of the same sign, the model is incomplete/incoherent with overlapping regions. See Figs. 1 and 2.

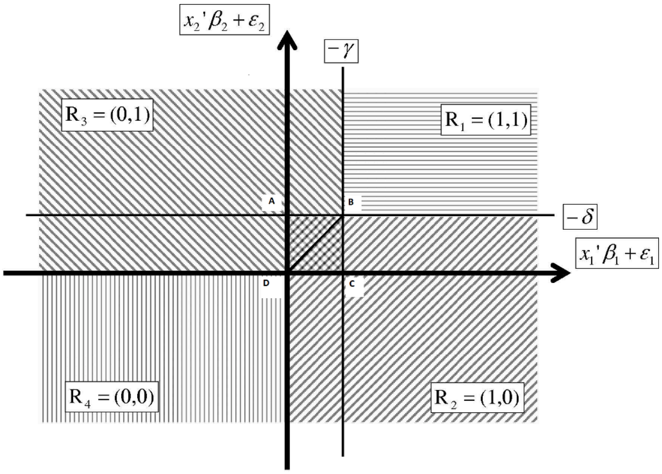

Fig. 2. Case 2: $\gamma < 0, \delta < 0$ Innovation and Finance Constraint bivariate binary probit in latent variables space. Four implied regimes. Region of Incoherency is of the overlap type (cross-hatched).

[4] Exceptions of this do exist, however: for example in industrial organization a two-agent discrete game may be employed to model the strategic interactions between firms in a duopoly setup. If one firm is a Stackelberg leader while the other is a follower, a recursive model may be applicable, even though the analogy is not precise.

[5] A related LDV model that does **not** exhibit similar coherency difficulties is the bivariate probit model with **latent variable** interactions (as opposed to limited variable interactions). Specifically:

$$y_{1it} = \tau_1\left(y^*_{1it} \equiv \left[h_1\left(x'_{1it}\beta_1, y^*_{2it}\gamma\right) + \epsilon_{1it}\right]\right)$$

$$y_{2it} = \tau_2\left(y^*_{2it} \equiv \left[h_2\left(x'_{2it}\beta_2, y^*_{1it}\delta\right) + \epsilon_{2it}\right]\right)$$

Then:

$$y_1^* = x_1\beta_1 + y_2^*\gamma + \epsilon_1$$

$$y_2^* = x_1\beta_1 + y_1^*\delta + \epsilon_2$$

and

$$y_1^* = x_1\beta_1 + \gamma \cdot \left[x_2\beta_2 + y_1^*\delta + \epsilon_2\right] + \epsilon_1$$

$$y_2^* = x_2\beta_2 + \delta \cdot \left[x_1\beta_1 + y_2^*\gamma + \epsilon_1\right] + \epsilon_2$$

Hence $y_1^* = RF_1$ and $y_2^* = RF_2$, allowing us to obtain $y_1 = \tau(RF_1)$ and $y_2 = \tau(RF_2)$. We thus see that it is considerably more straightforward identify LDV models with latent variable interactions as opposed to limited variable interactions.

[6] For example, economic theory suggests that $\gamma < 0$ (firms are less likely to be innovative if they cannot get external finance readily) and that $\delta > 0$ (banks are more likely to refuse loans to more "risky" innovative firms). The implied support restrictions are derived based on these economic considerations about the gamma and delta signs.

2.2. Cases 3 and 4: $\gamma \cdot \delta < 0$ — empty regions incoherency - identification through conditioning

When the interaction coefficients γ and δ are of opposite sign, the model exhibits empty regions incoherency. Identification can then be achieved by conditioning to lie outside the "empty" region. For example, Fig. 3 implies the conditioning probability: $1 - Prob(-\gamma < \epsilon_1 + x_1'\beta_1 < 0, 0 < \epsilon_2 + x_2'\beta_2 < -\delta)$. The estimation method that implements this is CMLE. The same procedure will solve the symmetric case depicted in Fig. 4 when $\gamma < 0, \delta > 0$.

2.3. Can overlapping regions incoherency be overcome through LDV redefinition?

In the absence of additional coherency conditions, there will be *overlaps* and/or *gaps* in the domain of $(\epsilon_1 + x_1'\beta_1, \epsilon_2 + x_2'\beta_2)$. A researcher might thus be tempted to propose that the incoherency cases with overlapping regions (Cases 1 and 2 above) may be overcome by redefining one of the two limited dependent variables to their complement. According to this reasoning, since the incoherency is caused in these cases because γ and δ are of the same sign, and since changing y_2, say, to its complement $y_2^N \equiv (1 - y_2)$ would result in $\delta^N \equiv -\delta$, then coherency would be achieved since then $\gamma \cdot \delta^N < 0$.

Such reasoning would be incorrect, however. We show in Appendix A that such a redefinition would *maintain* the overlapping-region incoherency or incompleteness. This is because the $y_2^N \equiv (1 - y_2)$ redefinition would also switch the sign of γ and hence $\gamma^N \cdot \delta^N > 0$ just as $\gamma \cdot \delta > 0$. See Fig. 5 for the analysis.

3. Simultaneously incomplete and incoherent models

The fundamental problem with the existing literature and with the model of the previous section is that it is not acknowledged that LDV Models can *simultaneously* exhibit both Incompleteness and Incoherency. We term such models *Simultaneously Incomplete and Incoherent (SII)* and illustrate with the following single-agent, simultaneous *binary & trinomial ordered probit model* of Hajivassiliou and Ioannides (2007), where a single optimizing agent responds rationally to multiple market constraints, such as financing/liquidity constraints and constraints on work hours. This model studies interactions between liquidity and employment constraints on individual households indexed by *i* at a given point in time indexed by *t*, and is in sharp contrast to the existing literature that concentrates on models where multiple agents play strategic games against each other. The reason we select this model is because it can exhibit simultaneously *both* types of incoherency (overlaps and gaps). This is critical because it will allow us to devise estimation strategies that overcome certain types of incoherency and incompleteness.

Define two latent dependent variables y_{1it}^* and y_{2it}^*. The first denotes the propensity of individual i in period t to be liquidity constrained and the second its propensity to face employment hour constraints. The corresponding LDVs are denoted by y_{1it} and y_{2it}. Dropping the *it* subscripts for simplicity, the model is defined by:

$$y_1 = \begin{cases} 1 & if \quad y_1^* > 0 \quad \text{(liquidity constraint binding)}, \\ 0 & if \quad y_1^* \leq 0 \quad \text{(liquidity constraint not binding)}. \end{cases} \tag{7}$$

$$y_2 = \begin{cases} -1 & if \quad y_2^* \leq \lambda^- \quad \text{(overemployed)} \\ 0 & if \quad \lambda^- \leq y_2^* < \lambda^+ \quad \text{(voluntarily employed)} \\ +1 & if \quad \lambda^+ \leq y_2^* \quad \text{(under-/unemployed)}. \end{cases} \tag{8}$$

where the latent variables are given by:

$$\begin{aligned} y_1^* &= \mathbf{1}(y_2^* < \lambda^-)\gamma_{11} + \mathbf{1}(\lambda^- < y_2^* < \lambda^+)\gamma_{12} + x_1'\beta_1 + \epsilon_1 \\ y_2^* &= \mathbf{1}(y_1^* > 0)\delta + x_2\beta_2 + \epsilon_2 \end{aligned}$$

Since (y_1, y_2) lie in $\{0,1\} \times \{-1,0,1\}$, the 6 possible configurations are:

y_1	y_2	y_1^*	y_2^*
0	-1	$\gamma_{11} + x_1\beta_1 + \epsilon_1 < 0,$	$x_2\beta_2 + \epsilon_2 < \lambda^-$
0	0	$x_1\beta_1 + \epsilon_1 < 0,$	$\lambda^- < x_2\beta_2 + \epsilon_2 < \lambda^+$
0	+1	$\gamma_{12} + x_1\beta_1 + \epsilon_1 < 0,$	$\lambda^+ < x_2\beta_2 + \epsilon_2$
1	-1	$\gamma_{11} + x_1\beta_1 + \epsilon_1 > 0,$	$\delta + x_2\beta_2 + \epsilon_2 < \lambda^-$
1	0	$x_1\beta_1 + \epsilon_1 > 0,$	$\lambda^- < \delta + x_2\beta_2 + \epsilon_2 < \lambda^+$
1	+1	$\gamma_{12} + x_1\beta_1 + \epsilon_1 > 0,$	$\lambda^+ < \delta + x_2\beta_2 + \epsilon_2$

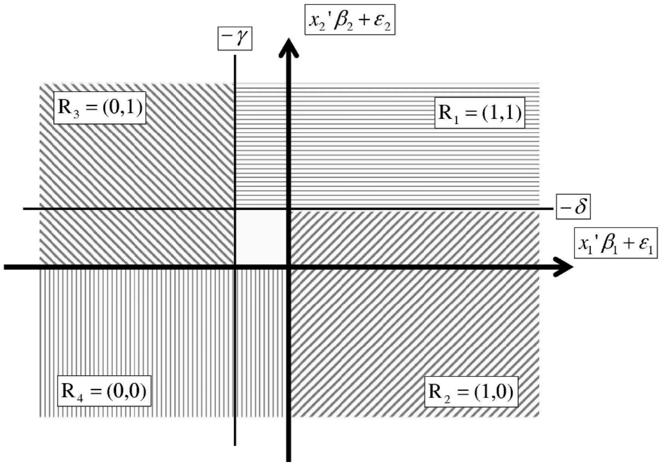

Fig. 3. Case 3: $\gamma>0,\delta<0$ Innovation and Finance Constraint bivariate binary probit in latent variables space. Four implied regimes. Region of Incoherency is of empty region type. Conditional MLE achieves coherency by ruling out the event: $[-\gamma<\epsilon_1+x_1'\beta_1<0, 0<\epsilon_2+x_2'\beta_2<-\delta]$

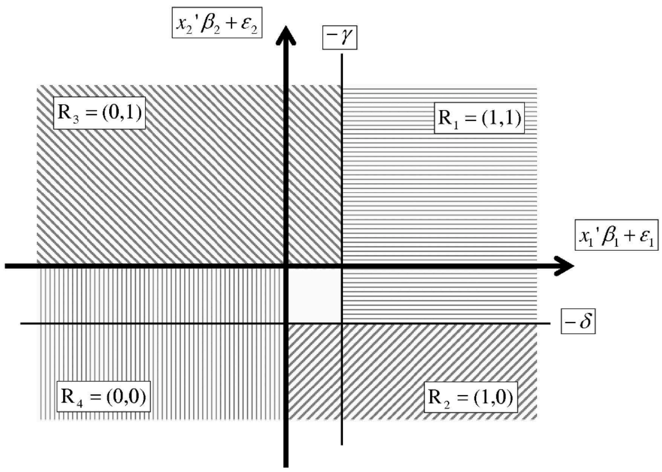

Fig. 4. Case 4: $\gamma<0,\delta>0$ Innovation and Finance Constraint bivariate binary probit in latent variables space. Four implied regimes. Region of Incoherency is of empty region type. Conditional MLE achieves coherency by ruling out the event: $[0<\epsilon_1+x_1'\beta_1<-\gamma, \delta<\epsilon_2+x_2'\beta_2<0]$

In terms of the unobservables, the probability of a (y_1,y_2) observed pair is equivalent to the probability: $\begin{pmatrix}a_1\\a_2\end{pmatrix}<\begin{pmatrix}\epsilon_1\\\epsilon_2\end{pmatrix}<\begin{pmatrix}b_1\\b_2\end{pmatrix}$ where $(\epsilon_1,\epsilon_2)'\sim N(0,\Sigma_\epsilon)$, and a and b are given by:

y_1	y_2	a_1	a_2	b_1	b_2
0	-1	$-\infty$	$-\infty$	$-(\gamma_{11}+x_1\beta_1)$	$\lambda^- - x_2\beta_2$
0	0	$-\infty$	$\lambda^- - x_2\beta_2$	$-x_1\beta_1$	$\lambda^+ - x_2\beta_2$
0	+1	$-\infty$	$\lambda^+ - x_2\beta_2$	$-(\gamma_{12}+x_1\beta_1)$	$+\infty$
1	-1	$-(\gamma_{11}+x_1\beta_1)$	$-\infty$	$+\infty$	$\lambda^- - \delta - x_2\beta_2$
1	0	$-x_1\beta_1$	$\lambda^- - \delta - x_2\beta_2$	$+\infty$	$\lambda^+ - \delta - x_2\beta_2$
1	+1	$-(\gamma_{12}+x_1\beta_1)$	$\lambda^+ - \delta - x_2\beta_2$	$+\infty$	$+\infty$

Using traditional arguments, we obtain that a sufficient condition for coherency of the model is: $(\gamma_{11}+\gamma_{12})\delta=0$ and $\gamma_{11}\gamma_{12}\delta=0$.

To verify this condition, suppose $(y_1,y_2)=(0,0)$. This rules out $(y_1,y_2)=(0,-1)$ because $x_2\beta_2+\epsilon_2>\lambda^-$, and rules out $(y_1,y_2)=(1,0)$ because $x_1\beta_1+\epsilon_1<0$. But $(1,-1)$ is not ruled out if the coherency conditions do not hold, since γ_{11} could be sufficiently negative and δ sufficiently positive to imply the $(1,-1)$ conditions. Similarly, the $(1,1)$ possibility cannot be ruled out in the absence of the coherency conditions, since γ_{12} and δ can be sufficiently positive.

Such logical inconsistencies are prevented if either (a) $\delta=0$ or (b) γ_{11} and γ_{12} are simultaneously 0.

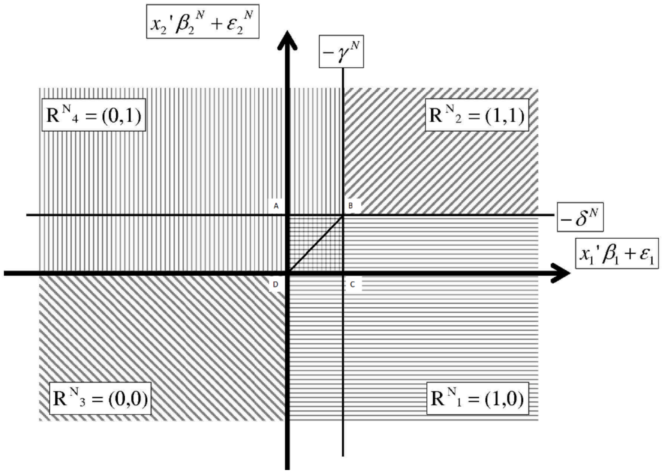

Fig. 5. $\gamma^N < 0, \delta^N < 0$ Innovation and Finance Constraint bivariate binary probit in latent variables space. Situation identical to Fig. 1 (Case 1) but with Innovation negated to its complement (1-Innovation). Both γ^N and δ^N switch sign by this redefinition, and hence: $\gamma^N < 0$, $\delta^N < 0$. Consequently, the model remains incoherent/incomplete, with an overlap (cross-hatched) region.

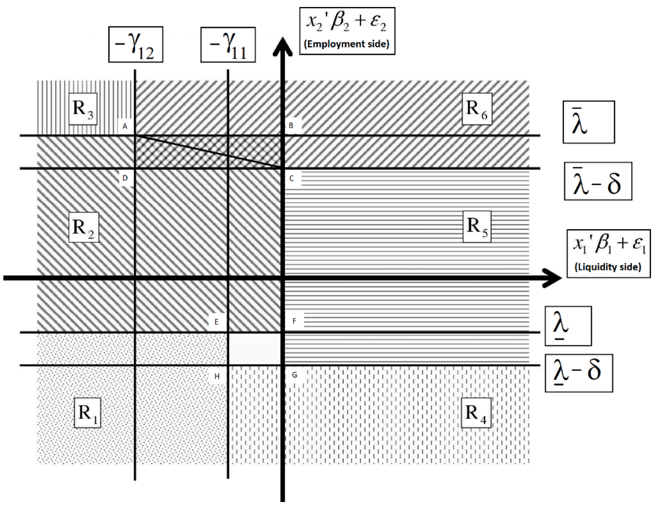

Fig. 6. Coherency and Completeness of Binary+Trinomial SII Model Regions in latent variables space that define the observed dependent qualitative variables Liquidity (binary) and Employment (trinomial) constraints. Six implied regimes. We see two regions of incoherency/incompleteness, one cross-hatched, the other empty. Both regions disappear under the sufficient condition: $\delta = 0$ and/or $\gamma_{11} = \gamma_{12} = 0$.

Fig. 6 gives the 6 possible regimes $(y_1 \times y_2) = \{1, 0\} \times \{-1, 0, 1\}$ in terms of the two latent variables y_1^* and y_2^* and the possible configurations in terms of parameters $\bar{\lambda} \equiv \lambda^+$, $\underline{\lambda} \equiv \lambda^-$, δ, γ_{11}, and γ_{12}. y_1^* is on the horizontal axis and y_2^* on the vertical.[7] The figure makes clear the role of the coherency condition (a) $\delta = 0$ or (b) $\gamma_{11} = \gamma_{12} = 0$: in general, regions R_2 and R_6 exhibit double-counting (cross-hatched area), as well as a white rectangle remains which makes the six regions not mutually exhaustive. These two logical incoherencies disappear when either $\delta = 0$ and/or $\gamma_{11} = \gamma_{12} = 0$ hold.

This graphical approach is useful to highlight the fundamental distinction between two types of incoherency, the first corresponding to overlap regions in latent variables space, while the second to empty regions. A critical fact that this model illustrates is that a particular model may **simultaneously** exhibit **incoherencies of both kinds,** *empty region incoherency* as well as *overlapping region incoherency,* also termed *incompleteness*. This is a critical point that is not well understood in previous work, e.g., Chesher and Rosen (2014).

[7] This graphical approach was first included in the LSE working paper Hajivassiliou (2008) and was presented at the CRETE Conference in Syros in 2003. It should also be noted that our graphical approach presented here is related to that of Tamer (2003) who studied the problem of coherency and incompleteness in bivariate discrete models for games with multiple equilibria.

4. Comparison with the existing literature

Heckman (1978) considered a simultaneous equations system with discrete endogenous regressors. He allowed both the empty region as well as the overlap region case in his proof of the main theorem based on his parameter π_{23}, whereby he showed that the necessary and sufficient condition (NASC) for the reduced form to "make sense" was the "*principal assumption*" that $\pi_{23} = 0$. He showed that if this coefficient is negative, the probabilities add up to below unity, whereas if positive they add up to over unity. Heckman did not introduce terms for either type of problem.

Gourieroux et al. (1980) explained the identification condition in terms of there being a valid function from $(\epsilon_{1it}, \epsilon_{2it})$ to the observable endogenous variables (y_{1it}, y_{2it}) so that empty region incoherency is ruled out. Maddala (1983) used the term "Logical Consistency" to mean the regime probabilities adding up to unity, but as with Heckman, the two problem cases were neither named nor discussed separately.

The first researcher to study overlap region incoherency, which he termed "Incompleteness," was Tamer (2003). It is useful to highlight here the similarities and differences with our work. Tamer also used a graphical approach to analyze a simultaneous discrete response model for a homogeneous two-agent discrete game of entry. Since the two rival firms in his setting were assumed identical, any incoherency arising was necessarily of the overlapping region type — see our two subcases 2.1, where the interaction terms γ and δ are of the same sign. Consequently, the possibility of the interaction terms being of opposite sign was not under focus in his analysis.

Subsequently, Lewbel (2007) derived the NASC for a valid reduced form to exist. However, the "Necessary" part implicitly ruled out the possibility of DGPs with *support restrictions*, which is precisely what underpins our CMLE solution.

It is useful to note that our approach of relying on prior sign restrictions developed here is related to the approach by Uhlig (2005) for Vector Autoregression identification under prior sign restrictions on impulse response functions.[8] Dagenais (1997) also makes a distinction between alternative types of incoherency regions.[9]

At this point, we stress that the approach and terminology in Chesher and Rosen (2014) is likely to lead to the incorrect belief that a model may exhibit only a single type of incoherency, either of the empty region- or overlap- type. Such misunderstanding would prevent the CMLE solution we develop in the following section. In fact, the terminology adopted by those authors, calling overlapping region incoherency as "model incompleteness" and empty region incoherency as proper "model incoherency," excacerbates this confusion in somehow giving the impression that a model can only exhibit one of the two irregularities.

It is also critical to explain why recent methodologies developed for econometric partially identified models with multiple equilibria *cannot* solve the coherency problems of the type we study here. There is a fundamental reason why the works of Berry and Tamer (2006), Ciliberto and Tamer (2009), Beresteanu et al. (2011) and de Paula and Tang (2012), which follow on the pioneering approach of Tamer (2003), are not applicable to our models: these works require simultaneous games with *multiple decision makers making a simultaneous decision*. In the absence of these two ingredients, we believe that our CMLE approach is the only available solution.

In addition, it is important to note the important work of Blundell and Smith (1994), which is an early example of Discrete/Continuous LDV models. In this paper we focus instead on purely Discrete simultaneous models, and outline our plans for investigating the more general Discrete/Continuous cases in Section 7.

Finally, we note that, contrary to Tamer (2003) and the cited subsequent game-theoretic papers, our DGP approach here can be extended to study the coherency of *dynamic* LDV econometric models. We develop these extensions in Section 6.1.

5. Coherency and completeness through parameter sign restrictions

The traditional approaches to model coherency suffer from several major difficulties. Firstly, derivations of formal conditions using the traditional approach lack intuition. Secondly, the derived conditions are impossible to generalize and verify in moderately more complicated LDV models, especially in cases where the models are allowed to contain intertemporal endogeneity of the type considered in Falcetti and Tudela (2008). Similarly, in case the joint binary probit model (3)-(4) is extended intertemporally, as for example in the empirical dynamic application in Section 8.2, the coherency condition is impossible to generalize and verify using the traditional analysis of the previous subsection. Thirdly, in practice non-triangular or reverse triangular cases are the most interesting from an economic point of view. Finally, the traditional approaches focus on establishing *sufficient* conditions for coherency, while our methods allow us to prove that they are *not necessary*.

To overcome the first two difficulties, alternative ways for establishing coherency are developed here, that are both intuitive and straightforward, as well as much more generalizable. In addition, our methods allow us to resolve the last two difficulties leading to estimation based on CMLE for much more interesting practical applications. It is shown in the next Section how to establish coherency without model recursiveness through the use of (a) endogeneity in terms of latent variables and/or (b) sign restrictions on model parameters.

[8] We are indebted to Alain Trognon for pointing out the potential of parameter sign restrictions overcoming incoherency of the "empty region" type, and to Hashem Pesaran for bringing to our attention Uhlig's work on sign identification.

[9] Unfortunately his work remains incomplete and unpublished due to his untimely death.

5.1. Efficient estimation through conditional maximum likelihood with empty region incoherency

The optimal parametric estimation approach for the models with empty region incoherency (Cases 3 and 4 above) will be *conditional maximum likelihood (CMLE)*, employing the appropriate likelihood contributions that characterize correctly the necessary conditioning through truncation that ensures that the LDVs stay out of the empty region of incoherency. For example, assuming independence across observations $\{it\}$, the likelihood contribution in Case 3 will be:

$$l_{it} = \frac{Prob(\epsilon_1, \epsilon_2 : I = 1(I^* > 0)\ \&\ F = 1(F^* > 0))}{(1 - Prob(-\gamma < \epsilon_1 + x_1'\beta_1 < 0, 0 < \epsilon_2 + x_2'\beta_2 < -\delta)}$$

while for Case 4:

$$l_{it} = \frac{Prob(\epsilon_1, \epsilon_2 : I = 1(I^* > 0)\ \&\ F = 1(F^* > 0))}{(1 - Prob(0 < \epsilon_1 + x_1'\beta_1 < -\gamma, \delta < \epsilon_2 + x_2'\beta_2 < 0)}$$

These likelihood contributions make it clear why approaches that ignore the coherency issue are inconsistent in general: the inconsistency would arise because the conditioning probability expressions in the denominator are functions of the underlying parameters and data, and hence affect critically the evaluation of the correct likelihood function.

It is important to remember that the fact that the likelihood contributions depend on denominator probabilities characterizing the support of the underlying truncated distributions that are also functions of parameters and data does *not* make the CMLE estimation problem irregular. The earliest example where such likelihood problems were studied formally is Amemiya (1973) for models of censoring and truncation. The uniform consistency, asymptotic normality, and efficiency of the CMLE estimators for the empty region inchorency Cases 3 and 4 can be established using methods in Amemiya (1973) and in works that followed.

Finally, we explain that the CMLE approach for achieving identification without triangularity and recursivity, does not violate the NASC results of Gourieroux et al. (1980) and Lewbel (2007) for a Valid Reduced Form to exist. This is because their "Necessary" part differs from ours in terms of the possible original DGP allowed, as they assume implicitly that the support of the original DGP is unrestricted. In contrast, our CMLE approach rests on the assumption that the true DGP is characterized by a truncated distribution so that the random draws stay out of incoherent regions.

5.2. Estimation with Overlap Region Incoherency

We note that Cases 1 and 2 (with same sign of the interaction coefficients γ and δ) may be handled in an analogous fashion *by assuming that* the Data Generating Process (DGP) incorporates additional mechanisms akin to the randomization postulated by Tamer (2003).[10] To find the correct likelihood contributions in Cases 1 and 2, define:

$$\begin{aligned} p_{11}^* &\equiv Prob(I^* > 0, F^* > 0) \quad p_{10}^* \equiv Prob(I^* > 0, F^* \le 0) \\ p_{01}^* &\equiv Prob(I^* \le 0, F^* > 0) \quad p_{00}^* \equiv Prob(I^* \le 0, F^* \le 0) \end{aligned}$$

Then, note that: $p_{11}^* + p_{10}^* + p_{01}^* + p_{00}^* = S = 1 + O > 1$ where O is the probability of the overlap region (rectangle ABCD in Figs. 1 and 2). In Case 1, the overlap occurs between regimes $R_1 = (1,1)$ and $R_4 = (0,0)$, while for Case 2 between regimes $R_2 = (1,0)$ and $R_3 = (0,1)$.[11]

In this paper, we propose two alternative additional identifying assumptions to complete the model in case of overlap region incoherency/incompleteness. The first approach, termed the "Contiguous Triangles," is to split the rectangle ABCD into the constituent triangles — ABC (closest to R_1) and ACD (closest to R_4) in Case 1, and BCD (to R_2) and ABD (to R_3) in Case 2 — and select the outcome of the triangle closest to the relevant regime. Thus, for Case 1 when the disturbances (ϵ_1, ϵ_2) fall in the overlap (indeterminate) region, triangle ABC is assigned to regime $R_1 = (1,1)$ and ACD to regime $R_4 = (0,0)$. Conversely, in Case 2, triangle BCD is assigned to regime $R_2 = (1,0)$ and triangle ABD to regime $R_3 = (0,1)$. In general, the probabilities of two equal triangles will differ depending on the shape of the density above them.

[10] The analogous estimation approach is presented here*for completeness only* because on a priori grounds the γ and δ of our Financing Constraints/Firm Innovation application are of opposite sign, a theoretical feature that is confirmed by our empirical findings. Furthermore, the randomization assumption is arguably much less unambiguous and clear-cut compared to the implied truncation in the cases of empty regions incoherency.

[11] Consequently, in Case 1 ($\beta > 0, \delta > 0$), $O = \Pr(-x_1\beta_1 - \gamma < \epsilon_1 < -x_1\beta_1, -x_2\beta_2 - \delta < \epsilon_2 < -x_2\beta_2)$, while in Case 2 ($\beta < 0, \delta < 0$) $O = \Pr(-x_1\beta_1 < \epsilon_1 < -x_1\beta_1 - \gamma, -x_2\beta_2 < \epsilon_2 < -x_2\beta_2 - \delta)$.

Consequently, the likelihood contribution for observation $\{it\}$ for Case 1 is:[12]

$$l_{it} = \begin{cases} p_{11} \equiv Prob(I = 1\&F = 1) = \left(p_{11}^{*} - \Pr(ACD)\right) \\ p_{10} \equiv Prob(I = 1\&F = 0) = p_{10}^{*} \\ p_{01} \equiv Prob(I = 0\&F = 1) = p_{01}^{*} \\ p_{00} \equiv Prob(I = 0\&F = 0) = \left(p_{00}^{*} - \Pr(ABC)\right) \end{cases}$$

A second alternative method we propose for overcoming *overlapping regions incoherency/incompleteness* is the "Truncated Distributions" approach. We then assume that the disturbances are drawn from distributions that are truncated to stay outside the indeterminacy/incompleteness rectangle ABCD. This enables CMLE, exactly analogously to the *empty region incoherency* cases. Given our emphasis on single-agent decision models, the two alternatives we propose here have a certain theoretical and behavioural appeal — whereas the alternative game-theoretic schemes[13] for the overlapping regions case are only applicable to models with multiple agents.

5.3. Estimation of SII models

As discussed above, for cases of overlapping regions incoherency/incompleteness, we proposed two alternative approaches for resolving the model: first, the "Contiguous Triangles" approach adjusts the overlap regime probabilities by assigning the two triangles of the overlap region ABCD to the closest regime. Second, the "Truncated Distributions" approach conditions on the probability of staying outside the indeterminancy region and is results in the CMLE parametrically optimal strategy. Both solutions preserve the correct individual regime probabilities and make them add up to unity. Also as explained already, in the cases of empty regions incoherency, the CMLE approach relies on truncating the DGP out of the empty regions, thus dividing the regime probabilities by the term $(1 - \Pr(\textit{empty region}))$.

In case of an SII model in which empty and overlap regions appear simultaneously, our proposed solutions can be combined. To illustrate with the 6-regime SII model (7)-(8) in Fig. 6, define $E = \Pr(\textit{empty region}) = \Pr(EFGH)$ and $O = \Pr(\textit{overlap region}) = \Pr(ABCD)$. Adopting the "Contiguous Triangles" incompleteness solution with the CMLE for the empty region incoherency, the adjusted probabilities are:

$$l_{it} = \begin{cases} p_1 \equiv Prob(S = 0\&E = -1) = p_1^{*}/(1 - E) \\ p_2 \equiv Prob(S = 0\&E = 0) = \left(p_2^{*} - \Pr(ABC)\right)/(1 - E) \\ p_3 \equiv Prob(S = 0\&E = +1) = p_3^{*}/(1 - E) \\ p_4 \equiv Prob(S = 1\&E = -1) = p_4^{*}/(1 - E) \\ p_5 \equiv Prob(S = 1\&E = 1) = p_5^{*}/(1 - E) \\ p_6 \equiv Prob(S = 1\&E = +1) = \left(p_6^{*} - \Pr(ADC)\right)/(1 - E) \end{cases}$$

After the adjustments, the probabilities add up to unity since:

$$\begin{aligned} \sum_{i=1}^{6} p_i &= \frac{p_1^{*} + \left(p_2^{*} - \Pr(ABC)\right) + p_3^{*} + p_4^{*} + p_5^{*} + \left(p_6^{*} - \Pr(ADC)\right)}{(1 - E)} \\ &= \frac{p_1^{*} + p_2^{*} + p_3^{*} + p_4^{*} + p_5^{*} + p_6^{*} - O}{(1 - E)} \\ &= \frac{1 - E + O - O}{1 - E} = 1 \end{aligned}$$

since $O = \Pr(ABCD) = \Pr(ABC) + \Pr(ADC)$. Identification then follows by the fact that the probabilities of the two types of regions, O and E, are distinct functions of the underlying parameters and the density function above them. Hence they will have distinct contributions to the likelihood derivatives.

[12] Analogously for Case 2:

$$l_{it} = \begin{cases} p_{11} \equiv Prob(I = 1\&F = 1) = p_{11}^{*} \\ p_{10} \equiv Prob(I = 1\&F = 0) = \left(p_{10}^{*} - Pr(ABD)\right) \\ p_{01} \equiv Prob(I = 0\&F = 1) = \left(p_{01}^{*} - Pr(BCD)\right) \\ p_{00} \equiv Prob(I = 0\&F = 0) = p_{00}^{*} \end{cases}$$

[13] (for example, the game-theoretic models of entry analyzed by Tamer (2003) and the works that followed him already cited)

6. Alternative analysis of incoherency and incompleteness: the DGP Method

Despite the usefulness of the graphical approach of the previous section to LDV problems with two latent variables, the method is very unwieldy or inapplicable to higher dimensional cases. To cover such problems, we develop a second approach to incoherency and incompleteness, which consists of designing a data-generating algorithm (hypothetical or implemented on a computer) to simulate random draws from an LDV model's structure. Consider again Liquidity-Employment Constraints Eqs. (7) and (8). We draw ϵ_1 and ϵ_2 under the joint bivariate normal distribution with zero mean vector and variance-covariance matrix Σ_ϵ, and given $x_1^{'}\beta_1$ and $x_2^{'}\beta_2$ we attempt to generate y_1^* and y_2^*. This is straightforward provided the coherency condition holds: If (a) $\delta = 0$, then latent y_2^* can be drawn, then LDV y_2, which together with ϵ_1 and $x_1^{'}\beta_1$ determines the right hand side of y_1^*, thus allowing y_1 to be drawn. Similarly, if (b) $\gamma_{11} = \gamma_{12} = 0$, then y_1^* can be drawn from the first equation based on ϵ_1 and $x_1^{'}\beta_1$, which determines y_1, thus giving y_2^* and hence y_2. In general it is impossible, however, to devise such a data generation mechanism in case the recursivity condition does not hold. Hence, the DGP method proves that recursity is sufficient to rule out empty-region and overlap-region incoherencies.

As we will show in the next section, the approach extends naturally to cases with intertemporal endogeneities in panel LDV models, and can be used to prove the coherency of the classic multiperiod panel probit with state dependence (Heckman (1981a)), as well as the intertemporal endogeneity versions of the models in Section 8.2 with explicit dynamic effects.

We show that empty region incoherency can be overcome through conditional maximum likelihood estimation (CMLE) of truncating the LDVs to lie outside the incoherency regions. The CMLE approach can also be motivated through the DGP approach for establishing coherency discussed in the previous subsection. In that case, we consider DGPs truncated to lie on a specific region of the latent variables space.[14]

Finally we show that under prior sign restrictions on model parameters, incoherencies of the empty region type can be eliminated through the use of CMLE.

6.1. *The coherency of dynamic LDV panel data (PD) models with intertemporal endogeneities using dgp approach*

Extending the analysis to a panel data set, we explain how the probability of a pair (y_{1it}, y_{2it}) in Section 2 and in Section 3, can be represented in terms of the linear inequality $(a_1, a_2)^{'} < (\epsilon_1, \epsilon_2)^{'} < (b_1, b_2)^{'}$where the error vector has a flexible autocorrelation structure. For example, one-factor random effect assumptions will imply an equicorrelated block structure on Σ_ϵ, while our most general assumption of one-factor random effects *combined with* an AR(1) process for each error implies that Σ_ϵ combines equicorrelated and Toeplitz-matrix features. Consequently, the approach incorporates fully (a) the contemporaneous correlations in ϵ_{it}, (b) the one-factor plus AR(1) serial correlations in ϵ_i, and (c) the dependency of y_{1it} on y_{2it} and vice versa. The coherency issue expands naturally to the panel sequence of data, by thinking of each (correlated) time-period for a given individual i as a **distinct probit equation** and then dealing with the independent cross-section of equations across individuals. Details of the analysis can be found in Hajivassiliou (2007).

The proofs of all Theorems and Lemmas can be found in Appendix B.

6.1.1. *Univariate PD Probit with State Dependence*

Our hypothetical DGP method presented above for establishing coherency is now applied to the canonical panel data Probit model with state dependence, first analyzed by Heckman (1981a). The model is defined by:

$$\begin{array}{rcl} y_{iT} & = & \mathbf{1}\big(\lambda y_{i,T-1} + x_{iT}\beta + \epsilon_{iT} > 0\big) \\ y_{i,T-1} & = & \mathbf{1}\big(\lambda y_{i,T-2} + x_{i,T-1}\beta + \epsilon_{i,T-1} > 0\big) \\ & & \vdots \\ y_{i2} & = & \mathbf{1}(\lambda y_{i1} + x_{i2}\beta + \epsilon_{i2} > 0) \\ y_{i1} & = & \mathbf{1}(x_{i1}\xi_1 + \cdots + x_{iT}\xi_T + u_{i1} > 0) \end{array}$$

The equation for $t = 1$ is a generalization of the Barghava and Sargan (1982) approach. Let $\Sigma \equiv VCov(\epsilon_{iT}, \cdots, \epsilon_{i1}, u_{i1})$. Imposing one-factor random effect assumptions will imply an equicorrelated block structure on the top left $(T-1) \times (T-1)$ block of Σ, while more general assumptions of one-factor random effects *combined with* an AR(1) or ARMA(p,q) processes for each ϵ error implies that Σ combines equicorrelated and Toeplitz-matrix parts. The last row and last column of Σ giving the variance of u_{1i} and its covariances with all ϵ_{it} allow the flexibility stipulated by Heckman (1981b).

Define the Cholesky lower triangular times upper triangular matrix factorization of $\Sigma = CC^{'}$. Given the assumed normality, the error vector can be written as: $(\epsilon_i^{'}, u_{1i})^{'} = C\nu_i$ where $\nu_i \sim N(0_T, I_T)$

Theorem 1. *The Univariate PD Probit Model with State Dependence defined above is coherent and complete.*

6.1.2. *Bivariate PD Probit with State Dependence*

Parameter mnemonics: (1) exogenous variable coefficients: β, θ; (2) interaction terms: γ, δ; (3) own state dependence: λ_y, λ_w; (4) cross state dependence: ζ_w, ζ_y.

[14] A specific algorithm for achieving this is given in Hajivassiliou (2008).

$$
\begin{aligned}
y_{iT} &= \mathbf{1}\left(x_{iT}^{'}\beta + \lambda_y y_{i,T-1} + \gamma w_{iT} + \zeta_w w_{i,T-1} + \epsilon_{iT} > 0\right) \\
w_{iT} &= \mathbf{1}\left(z_{iT}^{'}\theta + \lambda_w w_{i,T-1} + \delta y_{iT} + \zeta_y y_{i,T-1} + u_{iT} > 0\right) \\
&\cdot \\
y_{it} &= \mathbf{1}\left(x_{i,t}^{'}\beta + \lambda_y y_{i,t-1} + \gamma w_{it} + \zeta_w w_{i,t-1} + \epsilon_{it} > 0\right) \\
w_{it} &= \mathbf{1}\left(z_{it}^{'}\theta + \lambda_w w_{i,t-1} + \delta y_{it} + \zeta_y y_{i,t-1} + u_{it} > 0\right) \\
&\vdots \\
y_{i2} &= \mathbf{1}\left(x_{i2}^{'}\beta + \lambda_y y_{i1} + \gamma w_{i2} + \zeta_w w_{i1} + \epsilon_{i2} > 0\right) \\
w_{i2} &= \mathbf{1}\left(z_{i2}^{'}\theta + \lambda_w w_{i1} + \delta y_{i2} + \zeta_y y_{i1} + u_{i2} > 0\right) \\
&\cdot \\
&y_{i1}, w_{i1} \text{ exogenous}
\end{aligned}
$$

Lemma 2. *Without any restrictions on the γ, δ parameters or the distribution of (ϵ, u) , the General Bivariate PD Probit Model with State Dependence above is not coherent.*

Theorem 3. *The General Bivariate PD Probit Model with State Dependence above is coherent and complete without any restrictions on the λ, ζ state dependence parameters or the distribution of (ϵ, u), if the simultaneous interaction terms satisfy $\gamma \cdot \delta = 0$, i.e., the model is triangular.*

Theorem 4. *The General Bivariate PD Probit Model with State Dependence above is coherent and complete without any restrictions on the λ, ζstate dependence parameters, if:*

(i) the simultaneous interaction terms are of opposite signs, i.e., $\gamma \cdot \delta < 0$ and

(ii) the distribution of (ϵ, u) satisfies $F(\epsilon_t, u_t | \epsilon_{-t}, u_{-t}) = F(\epsilon_t, u_t | \epsilon_{<t}, u_{<t})$ and the random variables (ϵ, u) are restricted on rectangular regions determined recursively.

7. Issues and extensions

In Appendix C we develop and summarize the results of a set of extensive Monte-Carlo experiments, which confirm very substantive Mean-Squared-Error estimation improvements of the CMLE approach over estimators that make overly restrictive triangularity/recursivity assumptions about the Data Generating Process (DGP). We also discuss future extensions of the experiments to assess the performance of the proposed alternative “contiguous triangles” and “truncated distributions” solutions to the overlap incoherency/incompleteness problem.

Hajivassiliou (2007) discusses how to extend the analysis to another SII model with two simultaneous (bivariate) *ordered probit equations with multiple regions*. We refer the interested reader to that study.

Another natural extension to consider is to allow more complicated interaction terms rather than just the parallel shifts of the models of Section 2. For example, instead of the first Eq. (3), the extension would postulate:

$$
y_1 = \begin{cases} 1 & if \quad y_1^* \equiv x_1^{'}[\beta_1 \cdot y_2 + \beta_2 \cdot (1 - y_2)] + \epsilon_1 > 0 \\ 0 & if \quad y_1^* \equiv x_1^{'}[\beta_1 \cdot y_2 + \beta_2 \cdot (1 - y_2)] + \epsilon_1 \leq 0 \end{cases} \tag{9}
$$

This would allow possibly all the coefficients of the exogenous factors in the latent variables expressions to change depending on the binary outcome of the other side, instead of just the intercepts. This extension corresponds to a switching regressions specification with endogenous interaction terms.

An important question to pose is how important the exact distributions of the underlying disturbances are for our CMLE approach. The answer is that the distributional assumptions are important only to the extent that their supports are needed to characterize the necessary conditioning in order to achieve identification through CMLE, by ruling out ranges of values of the errors. They are also used for defining the necessary Accept-Reject truncation mechanism of the underlying DGP.

We plan to investigate reduced reliance on such parametric distributional assumptions, through semi-parametric simulation-based estimation of the probability of the empty regions (and possibly of overlap regions). This will be achieved by incorporating Accept-Reject simulation methods to define Conditional Maximum Simulated Likehood (CMSL) estimation approaches.

Another interesting extension is to focus on Discrete/Continuous LDV models, instead of the purely Discrete models of this paper — or the easier models with purely continuous interactions discussed in footnote 5. See, for example, the Roy models studied in, inter alia, Blundell and Smith (1994), Dubin and McFadden (1984), Hanemann (1984), and Heckman and Honoré (1990) where continuous endogenous variables are censored or truncated based on discrete simultaneous variables. Extending our analysis to such models is enhanced by certain features of the Roy model, while at the same time it is hindered by other features. A facilitating feature is the fact that the continuous side contains more valuable information than when purely discrete. On the other hand, the combination of continuous and discrete sides implies that the distributions that need to be analyzed are now much more complicated, namely censored, truncated, and/or mixtures of distributions. Along the same lines, the final extension we intend to study are Discrete/Continous Endogenous Treatment models, whereby outcome Y depends on treatment T and simultaneously T depends on Y. We leave these extensions to future work.

Table 1
Comparative Summary of Empirical Results

All Firms	Model 1 Triangular, Exogenous FC	Model 2 Triangular, Endogenous FC	Model 3 Full Joint Static	Model 4 Full Joint Dynamic
No.of Waves	One	One	One	Two
No.of Firms	1940	1940	1940	1512
INNOVATION EQUATION				
Size	0.33***	0.305	0.183***	0.256***
γ (FC dummy)	0.55***	-0.555***	-0.324**	-0.447***
Market Share	-0.01	-0.001***	0.020	0.027
Innov$_{t-1}$	—	—	—	0.829***
FinCon$_{t-1}$	—	—	—	0.301
avg $\widehat{P_I}$	0.418	0.418	0.418	0.543
$\widehat{P_I}: F = 0$	0.384	0.453	0.438	0.554
$\widehat{P_I}: F = 1$	0.601	0.250	0.316	0.377
$\%\widehat{\Delta P_I}: F = 0\rightarrow 1$	56.42	-44.72	-27.93	-31.82
FINANCE EQUATION				
Size	-0.054	-0.002	-0.016	0.035
δ (Innov dummy)	—	—	0.647***	0.627***
Collateral	0.067	0.030	0.030	0.003
Banking Debt Ratio	0.010***	0.010***	0.015***	0.005
Own Financing Ratio	-0.003**	-0.003***	-0.001**	-0.008***
Profit Margin	-0.007**	-0.008***	-0.002***	-0.007***
Innov$_{t-1}$	—	—	—	0.236**
FinCon$_{t-1}$	—	—	—	0.135***
avg $\widehat{P_F}$	0.160	0.160	0.160	0.060
$\widehat{P_F}: I = 0$	0.160	0.160	0.103	0.029
$\widehat{P_F}: I = 1$	0.160	0.160	0.268	0.102
$\%\widehat{\Delta P_F}: I = 0\rightarrow 1$	0.0	0.0	160.62	252.58
corr(Innov,FinCons)	—	0.572***	0.132**	0.500**
LogLikFunction	-1060-803=-1863	-1853	-1712	-1331 (-1706 imputed)

1. ***=significant at 1%; **=significant at 5%; *=significant at 10%.2. Industry dummies (11) included in both Innovation and Financial Constraint equations.

8. Quantifying the interactions between financial constraints and firm innovation

8.1. *Direct measures of innovation and financial constraints*

We apply our approach to assess the links between financing constraints and a firm's decision to innovate. Empirical evidence of the impact of financial constraints on the behaviour of firms is however not easy to obtain, essentially because the notional demand of firms for external finance is not observed directly.[15] We rely on qualitative direct information about innovation and binding financing constraints as reported by firms in surveys by the European Union (CIS, Community Innovation Surveys), as well as in a French survey about the financing of innovation (FIT, Financement de l'innovation technlogique). We merged the surveys with balance sheet dataset (Fiben, Banque de France), see Data Online Appendix for details.[16]

In the survey we use, firms are asked whether some of their innovative projects were delayed, abandoned or non started because of (i) unavailability of new financing, (ii) searching and waiting for new financing, (iii) too high cost of finance. We define as financially constrained firms with hampered innovative projects because of one of these three reasons, so that our direct indicator of financial constraints takes into account both quantity rationing and higher cost of finance.[17] Innovation in this survey is defined according to the technological innovation in the Oslo Manual (OECD).[18]

[15] See Hottenrott and Peters (2012) for a test based on the use of a hypothetical payment received by the firm, and Brown et al. (2009), Brown et al. (2012) for traditional approaches linking R&D investment and firm level financial factors.

[16] Such direct information about credit constraints have already been used by Guiso (1998) to assess the probability for a firm to be credit constrained, and more recently by Minetti et al. (2019) to study the effect of financial constraints on firms' participation in supply chains.

[17] Given that this is a qualitative self-assessed measure of financial constraint, we checked that it correlates strongly with quantitative balance sheet variables related to the financial health of firms. (See Table A1 and Table A2 in the online data appendix for summary descriptive statistics).

[18] This is then a qualitative self-assessed and survey-based information which was implemented to overcome some other shortcomings of traditional measures (R&D, patents), see Mohnen (2019) for a detailed discussion on these issues.

8.2. Empirical application

Using the econometric machinery developed in Section 5.1 that allows us to estimate joint binary probit equations with interaction terms on both sides, we apply those methods to the key issue of Being Innovative vs. Binding Financing Constraints interactions. Obviously, the propensity to innovate may be affected by financial constraints, and at the same time, innovative firms are likely to face specific financial constraints: innovation affects survival of firms (see Audretsch (1995) and Klette and Kortum (2004)), asset intangibility is higher for innovative firms which lowers their collateral value, and due to their innovative nature informational asymmetries with external investors are more pronounced (Holmstrom (1989), Hall and Lerner (2010)).

We take as our starting point the results obtained by Savignac (2008) who studied the impact of financial constraints on the decision to innovate by investigating the impact of financial constraints on innovation through a recursive model that did not allow for the probability of a binding finance constraint to depend on whether or not the firm is innovative. Recursivity corresponds to the key identifying assumption that innovation does not affect financial distress directly ($\delta = 0$). On a priori grounds, this assumption seems particularly dubious since innovation may lead to higher profits and thus relax financial constraints (corresponding to $\delta > 0$). An alternative possibility is that innovation may lead to higher investment in intangible assets thus reinforcing binding financial constraints (corresponding to $\delta < 0$). Both possibilities violate the traditional coherency condition.

The propensity to innovate is explained by the traditional determinants of innovation, and we account for financial constraints through our qualitative indicator reflecting the financial difficulties encountered by firms to conduct their innovative projects.[19] To close the system, we now define also the probability of a binding financing constraint, which is assumed to have as an important determinant the (binary) decision of whether or not the firm chooses to be innovative. The key idea modelled by this equation is that prospective lenders will try to assess the creditworthiness of the applicant firm in the face of incomplete information. In particular, they do not know the precise riskiness of assets so they attempt to infer that using observable characteristics of firms. In the face of such uncertainty, it makes sense for lenders to be more cautious granting loans to innovative firms since they present a higher inherent (but not directly observed) risk.

Such a system can be formulated as in Eqs. 3-(4) of Section (2), where x_i^I is a list of control variables for the traditional determinants of the propensity to innovate (firm size and market power, technology push, latent consumer demand);[20] and x_i^F is a list of control variables for the propensity to be finance constrained (firm's size, profit margins, the guarantees that could be used as collateral, industry dummies, and variables reflecting the firm's ex ante financial structure). See the Online Appendix for the detailed list of control variables.

The econometric specifications we estimate below belong in three main groups. The first group contains *recursive specifications* following the traditional approach, which ignore the possibility that the propensity to innovate may be affected by financial constraints. The second group allows for *reverse interactions*, whereby a firm undertaking actively innovative activities raises significantly the probability of it encountering a binding financing constraint, possibly because potential lenders are particularly wary of granting loans to firms of such type because of the extra riskiness involved. The third group of estimated specifications, investigates *state dependence* in financing and innovation experiences of firms the nature of the available datasets can be exploited to study whether, ceteris paribus, past financial distress or innovation failures can affect a firm's current experiences in these two dimensions.[21] The second and third groups of specifications constitute the key applications of the novel estimation strategies developed in this paper.

The most general specification that we estimate below, accounting for both reverse and dynamic effects, is:

$$I_{it} = \begin{cases} 1 & if \quad I_{it}^* \equiv \alpha^I I_{it-1} + x_{it}^I \beta^I + \gamma_0 F_{it} + \gamma_1 F_{it-1} + \epsilon_{it}^I > 0 \\ 0 & if \quad I_{it}^* \equiv \alpha^I I_{it-1} + x_{it}^I \beta^I + \gamma_0 F_{it} + \gamma_1 F_{it-1} + \epsilon_{it}^I \le 0 \end{cases} \tag{10}$$

$$F_{it} = \begin{cases} 1 & if \quad F_{it}^* \equiv \alpha^F F_{it-1} + x_{it}^{F'} \beta^F + \delta_0 I_{it} + \delta_1 I_{it-1} + \epsilon_{it}^F > 0 \\ 0 & if \quad F_{it}^* \equiv \alpha^F F_{it-1} + x_{it}^{F'} \beta^F + \delta_0 I_{it} + \delta_1 I_{it-1} + \epsilon_{it}^F \le 0 \end{cases} \tag{11}$$

[19] For the importance of endogeneity in this setting, see Mohnen and Roller (2005) who find the "paradox" of a positive correlation between financial constraints and innovation.

[20] Main determinants of the propensity for a firm to innovate are known to be its size, its market power and its environment ([Cohen and Levin, 1989]). The positive correlation between innovation and firm size is largely exposed in the literature. (see [Cohen and Klepper, 1996]). Large firms can amortize sunk costs caused by their innovative activities and are able to diversify the risk incurred by innovation by running simultaneously several investment projects at the same time. And finally, large established firms are less likely to be financially constrained as they are able to generate cash-flow and to raise external funds. Regarding the link between innovation and competition, the Schumpeterian theory argues that market power and innovation are positively correlated whereas Arrow's theory shows that the gains to innovate are larger in an ex-ante competitive market. [Aghion et al., 2005] try to solve this puzzle and propose an inverted U shape relationship between innovation and competition: in a competitive environment, firms are incited to innovate to gain market power and increase their profits, but when competition becomes hard, the followers can be discouraged to innovate. Other factors affecting innovative behaviour are driven by the firm environment (technological push, latent consumer demand perceived by the firm). See among others [Crepon et al., 1998] or [Raymond et al., 2010].

[21] Though the surveys about innovation we use are not truly longitudinal "panel" sets, the information was collected in multiple biennial waves. We hence restrict our "longitudinal" dataset to two waves in order to limit the reduction of the sample size when merging the waves. Hence, we know whether a particular firm *i* has reported binding financing constraints in the past. See the Online Data Appendix for details about the dataset employed here and the transition tables for the 1512 firms in the sample over the two waves starting 1994 and ending 1999.

8.3. Comparative analysis of alternative specifications

Table 1 summarizes succinctly our key empirical results and presents the calculated direct and reverse effects that we obtained.[22] To recapitulate, Models 1 and 2 adopt the existing approach of the past literature of forcing the econometric specification to be triangular with financial constraints allowed to affect the innovation decision, while the financial constraint outcome is assumed independent of being innovative or not. Models 3 and 4 allow the binary interactions on both the Innovation and the Finance sides, proved to be coherent through our Coherency analysis based on prior sign restrictions and estimated through the CMLE approach developed and analyzed in this paper. Model 3 is static with a single cross-section of firms, while Model 4 uses a two-wave, dynamic panel data set.

Apart from the coefficient estimates for the most important exogenous explanatory variables, Table 1 presents also: γ, the coefficient for the financial constraint dummy when entered in the Innovation side; δ, the coefficient for the Innovation dummy entered in the Finance side; and for the dynamic Model 4, the coefficients for the lags of the Finance and Innovation dummies. First, note how seriously misleading conclusions were reached by the early strands of the literature, that inappropriately ignored the endogeneity of the Finance dummy: doing so yields a completely counterintuitive γ estimate of $+0.55$, implying that finance constraints *raise* significantly the likelihood of innovation, confirming Savignac (2008). This positive effect is explained by two sources of bias: a selection bias due to firms not wishing to innovate, which we studied elsewhere and a problem of simultaneity between investment and financing decisions that we tackle below. At the same time, treating the Finance constraint dummy as endogenous gives a range of negative estimates from -0.32 to -0.56. Since Models 1 and 2 impose a triangular specification, they do not estimate δ coefficients for the Innovation dummy in the Finance equation. In contrast, our two CMLE models give statistically very significant δ estimates of over 0.6 — as we expected a priori, being innovative raises the probability of facing a binding finance constraint. In the dynamic Model 4, we find very significant state dependence over the two periods of the panel — note the statistical significance of three of the four lagged dummies entered as regressors. Our findings confirm the strong importance of such dynamic terms and establish very significant positive state dependency. Our results show that firms tend to innovate persistently rather than occasionally.

Second, past financial difficulties are positively correlated with current binding financial constraints. As we take into account the experience of a firm concerning innovation, the state dependence of financial constraints seems particularly interesting. Indeed, firms currently implementing innovative projects as well as firms with innovative experience in the past are more likely to find it difficult to finance their current projects.[23]

Third, the probability for a firm to be currently conducting an innovative project is negatively impacted by the current financing difficulties as found in the static regressions but also positively correlated with financing constraints encountered in the past. One possible explanation for this positive correlation could be that financial difficulties mainly impact the beginning of the projects so that innovative projects that were initially hampered by financial difficulties are more likely to be continued when they become more mature. However, additional information on the stage of development of the innovative projects and on their duration would be necessary to further investigate this point. In particular, we are not able to identify whether the firms were continuing with the same project over Waves 1 and 2.

In order to quantify the importance of our interaction findings about γ and δ, we present in Table 1 four estimated probability calculations for each of the two I and F sides: (a) *avg* $\widehat{P_I}$, the average probability of undertaking innovation; (b) $\widehat{P_I} : F = 0$, the estimated probability of Innovation given the Finance constraint is not binding; (c) $\widehat{P_I} : F = 1$, the estimated probability of Innovation given a binding Finance constraint; and (d) $\%\widehat{\Delta P_I} : F = 0 \rightarrow 1$, the percentage change in the estimated probability of changing from $F = 0$ to $F = 1$, while keeping everything else unchanged. For the finance side, the analogous four quantities are: *avg* $\widehat{P_F}$, $\widehat{P_F}: I = 0$, $\widehat{P_F}: I = 1$, and $\%\widehat{\Delta P_F}: I = 0 \rightarrow 1$.

Our estimated probability results are quite striking: for the Innovation equation, we find that *when a firm faces a binding finance constraint, the probability of being innovative falls ceteris paribus by 30%-40%* depending on the version.[24] Moving to the Finance constraint side, the magnitudes of the results are even more impressive: *a firm being innovative more than doubles the probability of a finance constraint.*[25]

[22] The detailed estimation results are presented in Tables 2 and 3 in Appendix D.

[23] An important issue discussed frequently in the econometrics literature is the possibility that state dependence may not be an important factor *per se*, but it might appear statistically significant if persistent heterogeneity among individual economic agents is ignored. As Heckman (1981a) shows, the two can be identified when a panel data set with more than two waves per individual is available. Since our dynamic sample consists only of two waves, we need to acknowledge the possibility that the strong state dependence we report here may be compounded by unobserved persistent heterogeneity that is not accounted for explicitly.

[24] In the erroneous Model 1 that ignores the endogeneity of the Finance constraint, the probability of innovation is predicted to*rise by over* 50% as a result of a binding Finance constraint!

[25] Since Models 1 and 2 do not allow for reverse interactions by excluding the I dummy from the F side, they imply $\%\widehat{\Delta P_F}: I = 0 \rightarrow 1$ equal to zero.

9. Conclusions

We investigated the fundamental identification issues of *coherency* and *completeness* of LDV models with endogeneity and flexible temporal and contemporaneous correlations in the unobservables. We focused on the class of models that can be Simultaneously Incoherent and Incomplete (SII), which was largely ignored by the existing literature. We defined and distinguished two types of identification problems termed "empty-region incoherency" and "completeness or overlap-region incoherency." The distinction cleared up the past literature where confusing and frequently conflicting definitions were used. The class of SII models establishes that incoherency and incompleteness are not a model's either/or properties but they could co-exist. Furthermore, while the traditional approaches focus on *sufficient* conditions for coherency, our methods prove that they *may not be necessary*.

We proceeded to establish coherency and completeness of static and dynamic LDV models without imposing recursivity. We achieved this through the use of sign restrictions on model parameters. The issues and solutions were first presented through a Graphical approach, while we then developed an alternative analysis of incoherency and incompleteness which we term the DGP approach. Using the DGP approach we solved considerably more complicated LDV models, especially those containing intertemporal endogeneity such as the Panel Univariate and Panel Bivariate Probit models with State Dependence.

The two novel methods we developed led to CMLE for simultaneous LDV models and allowed us to establish the coherency and completeness of several Dynamic LDV models which was impossible using traditional methods.

We proposed an empirical application to quantify the interactions between innovation by firms and the financial constraints they may face. We used as the modelling framework the Panel Bivariate Probit model with State Dependence.

Our empirical findings were quite striking: ceteris paribus, we estimated that a firm that faces a binding finance constraint is approximately *30%-40% less likely* to undertake innovation, while the probability that a firm encounters a binding finance constraint *more than doubles* if the firm is classified as innovative. Finally, we established a strong role for state dependence: firms tend to innovate continuously rather than occasionally; past financial difficulties are correlated with the present ones even after conditioning on important firm characteristics; and firms with current and/or past innovative experiences are more likely to encounter difficulty financing their current projects. Our novel approach for the first time eliminated the need to assume model recursivity. We concluded that such issues are critical if direct and reverse interactions between innovation and financing constraints are to be quantified reliably.

Appendix A: Redefining Binary Response Indicators

Suppose we have incoherency because we believe $\gamma > 0$ (in our application below translating to binding finance constraints expected to raise the chance of innovation *I*) and that $\delta > 0$ (innovative firms face a higher chance that the banks will refuse them a loan). So $\gamma \cdot \delta > 0$. This is Case 1 analyzed in Subsection 2.1 as represented by Fig. 2, and corresponding to the constraints on the unobservables: $(a^1, a^2)' < (\epsilon^1, \epsilon^2)' < (b^1, b^2)'$ such that:

I	F	a^1	b^1	a^2	b^2	Shading	Region
1	1	$-x_1'\beta_1 - \gamma$	∞	$-x_2'\beta_2 - \delta$	∞	horizontal	R1
1	0	$-x_1'\beta_1$	∞	$-\infty$	$-x_2'\beta_2 - \delta$	///////	R2
0	1	$-\infty$	$-x_1'\beta_1 - \gamma$	$-x_2'\beta_2$	∞	\\\\\\\\\\\\\\	R3
0	0	$-\infty$	$-x_1'\beta_1$	$-\infty$	$-x_2'\beta_2$	vertical	R4

Now consider the transformed model with *NF* **instead of** *F*. This transformation still gives an overlapping region in the transformed variables, and hence corresponds to an incoherent model. To see this, proceed as follows:

In terms of the two latent variables I^* and $NF^* = -F^*$ and the observed binary indicators *I* and $NF = 1 - F$, and suppressing the observation index:

$$I = \begin{cases} 1 & if \quad I^* \equiv x_1'\beta_1 + \gamma^N NF + \epsilon_1 > 0 \\ 0 & if \quad I^* \equiv x_1'\beta_1 + \gamma^N NF + \epsilon_1 \le 0 \end{cases} \tag{13}$$

$$NF = \begin{cases} 1 & if \quad NF^* \equiv x_2'\beta_2^N + \delta^N I + \epsilon_2^N > 0 \\ 0 & if \quad NF^* \equiv x_2'\beta_2^N + \delta^N I + \epsilon_2^N \le 0 \end{cases} \tag{14}$$

Given this transformation, we expect that $\gamma^N < 0$ (high *NF* means not very binding constraints so cause dampening of *I*) and that $\delta^N < 0$ (firms who have high *I* i.e., innovate, raise the chance the banks will refuse them a loan so low *NF*). So $\gamma^N \cdot \delta^N > 0$. See Fig. 5.

For a typical *i* observation, the probability $Prob(y_{1i}, y_{2i}|X, \theta)$ is characterized by the constraints on the unobservables: $(a^1, a^2)' <$

$(\epsilon_1, \epsilon_2^N)' < (b^1, b^2)'$ through the configuration:

I	NF	a^1	b^1	a^2	b^2	Shading	Region
1	0	$-x_1'\beta_1$	∞	$-\infty$	$-x_2'^N\beta_2 - \delta^N$	horizontal	R1
1	1	$-x_1'\beta_1 - \gamma^N$	∞	$-x_2'\beta_2^N - \delta^N$	∞	///////	R2
0	0	$-\infty$	$-x_1'\beta_1$	$-\infty$	$-x_2'^N\beta_2$	\\\\\\\\\\\\\\	R3
0	1	$-\infty$	$-x_1'\beta_1 - \gamma^N$	$-x_2'^N\beta_2$	∞	vertical	R4

Appendix B: Proofs of Theorems

Proof of Theorem 1: We use the DGP approach. Let us begin with the simplified case of the initial condition being exogenous:

$$y_{iT} = \mathbf{1}\left(\lambda y_{i,T-1} + x_{iT}'\beta + \epsilon_{iT} > 0\right) \tag{14}$$

$$y_{i,T-1} = \mathbf{1}\left(\lambda y_{i,T-2} + x_{i,T-1}'\beta + \epsilon_{i,T-1} > 0\right) \tag{15}$$

$$\vdots \tag{16}$$

$$y_{i2} = \mathbf{1}\left(\lambda y_{i1} + x_{i2}'\beta + \epsilon_{i2} > 0\right) \tag{17}$$

$$y_{i1} = \textbf{exogenous} \tag{18}$$

Suppose first the ϵ_{it} has the one-factor (equicorrelated) error components structure $\epsilon_{it} = \alpha_i + \nu_{it}$. Conditional on α_i, these $T - 1$ equations are independent (since they only depend on the i.i.d. ν_{it}s). Hence draw an α_i and an independent ν_{i2}. Then use the exogenous y_{i1} outcome to generate y_{i2}. This completes Eq. 17 which allows to move sequentially to generating y_{i3}, then y_{i4}, etc. until y_{iT} is generated. This establishes the coherency of the model.

Now allow for a general $\Sigma \equiv VCov(\epsilon_{iT}, \cdots, \epsilon_{i2}) = CC'$. Given that we assume Gaussianity and dropping the i index, we obtain:

$$y_T = \mathbf{1}\left(\lambda y_{T-1} + x_T\beta + c_{T1}\nu_1 + c_{T2}\nu_2 + \cdots + c_{T,T-1}\nu_{T-1} + c_{TT}\nu_T > 0\right)$$

$$y_{T-1} = \mathbf{1}\left(\lambda y_{T-2} + x_{T-1}\beta + c_{T-1,1}\nu_1 + c_{T-1,2}\nu_2 + \cdots + c_{T,T-1}\nu_{T-1} > 0\right)$$

$$\vdots$$

$$y_2 = \mathbf{1}(\lambda y_1 + x_2\beta + c_{22}\nu_2 + c_{21}\nu_1 > 0)$$

$$y_1 = \textbf{exogenous}$$

Given a random draw of $\nu_{i1}, \cdots, \nu_{iT}$, an unambiguous rule gives sequentially $y_{i1} \rightarrow y_{i2} \rightarrow \cdots y_{i,T-1} \rightarrow y_{iT}$. Hence, the above defines a recursive DGP which establishes the coherency of the model.

Finally, consider the more general case when y_{i1} cannot be assumed as exogenous. We then supplement the system with an initial condition equation:

$$y_{i1} = \mathbf{1}(x_{i1}\xi_1 + \cdots + x_{iT}\xi_T + u_{i1}))0 \tag{19}$$

The following remarks are in order: First note that (19) is a generalization of the Barghava and Sargan (1982) approach. Second, one-factor random effect assumptions will imply an equicorrelated block structure on the top left $(T-1) \times (T-1)$ block of Σ, while more general assumptions of one-factor random effects *combined with* an AR(1) or ARMA(p,q) processes for each ϵ error implies that Σ combines equicorrelated and Toeplitz-matrix parts. The last row and last column of Σ giving the variance of u_{1i} and its covariances with all ϵ_{it} allow the flexibility stipulated by Heckman (1981a). The only modification now necessary is to change the initial condition equation to:

$$y_{i1} = \mathbf{1}(x_{i1}\xi_1 + \cdots + x_{iT}\xi_T + c_{11}\nu_{i1} > 0)$$

This recursive representation again establishes the coherency of the model: given a random draw of $\nu_{i1}, \cdots, \nu_{iT}$, an unambiguous DGP rule can be defined to establish sequentially $y_{i1} \rightarrow y_{i2} \rightarrow \cdots y_{i,T-1} \rightarrow y_{iT}$. ■

Proof of Lemma 1:

$$\begin{aligned} y_{it} &= \mathbf{1}\left(x_{it}^{'}\beta + \lambda_y y_{i,t-1} + \gamma w_{it} + \zeta_w w_{i,t-1} + \epsilon_{it} > 0\right) \\ w_{it} &= \mathbf{1}\left(z_{it}^{'}\theta + \lambda_w w_{i,t-1} + \delta y_{it} + \zeta_y y_{i,t-1} + u_{it} > 0\right) \end{aligned}$$

Given y, w from period $t-1$, the λ and ζ terms are determined on the latent variable terms for period t (defining the event arguments of the indicator functions).

Together with unrestricted values of the random shocks and the exogenous variables of period t, everything in the event conditions is determined, except the simultaneous interaction terms γ, δ.

But since the interaction terms appear both as conditioning variables on the RHS as well as dependent variable dummies on the LHS, they cannot be determined unambiguously. Hence, no complete DGP can be defined from ϵ, u to y, w. ■

Proof of Theorem 2: Assume that $\gamma \cdot \delta = 0$ because $\gamma = 0$.

$$\begin{aligned} y_{it} &= \mathbf{1}\left(x_{it}^{'}\beta + \lambda_y y_{i,t-1} + \zeta_w w_{i,t-1} + \epsilon_{it} > 0\right) \\ w_{it} &= \mathbf{1}\left(z_{it}^{'}\theta + \lambda_w w_{i,t-1} + \delta y_{it} + \zeta_y y_{i,t-1} + u_{it} > 0\right) \end{aligned}$$

Given y, w from period $t-1$, the λ and ζ terms are determined on the latent variable terms for period t (defining the event arguments of the indicator functions). Together with unrestricted values of the random shocks and the exogenous variables of period t, everything in the event condition of the y_t is determined, since no simultaneous interaction term is present on the RHS (as $\gamma = 0$).

Entering the y_t value in the interaction term on the RHS of the w_t equation, everything in its event condition is now determined, which fixes w_t.

Hence, a complete DGP can be defined sequentially from the errors to the observables: $y_{i1}, w_{i1} \rightarrow y_{i2}, w_{i2} \rightarrow \cdots y_{i,T-1}, w_{i,T1} \rightarrow y_{iT}, w_{iT}$.

The proof for the $\delta = 0$ case is perfectly symmetric and will not be repeated. ■

Proof of Theorem 3: Assume that $\gamma \cdot \delta < 0$ because $\gamma < 0, \delta > 0$.

$$\begin{aligned} y_{it} &= \mathbf{1}\left(x_{it}^{'}\beta + \lambda_y y_{i,t-1} + \gamma w_{it} + \zeta_w w_{i,t-1} + \epsilon_{it} > 0\right) \\ w_{it} &= \mathbf{1}\left(z_{it}^{'}\theta + \lambda_w w_{i,t-1} + \delta y_{it} + \zeta_y y_{i,t-1} + u_{it} > 0\right) \end{aligned}$$

Given y, w from period $t-1$, the λ and ζ terms are determined on the latent variable terms for period t (defining the event arguments of the indicator functions).

Given the exogenous variables of period t, the event conditions of y_t, w_t are determined except (a) the interaction terms γ, δ and (b) the error terms.

In the absence of condition (ii), the model would exhibit "empty region incoherency" as defined above. Employing the graphical approach of the static bivariate probit above, defines the necessary rectangular exclusion region (drawn white) for the support of the truncated gaussian:

$$\begin{aligned} 0 &< \epsilon_{it} + x_{it}^{'}\beta_t + \lambda_y y_{i,t-1} + \zeta_w w_{i,t-1} < -\gamma \\ \delta &< u_{it} + z_{it}^{'}\theta + \lambda_w w_{i,t-1} + \zeta_y y_{i,t-1} < 0 \end{aligned}$$

Based on the underlying uniform rv's drawn at the start of the DGP, the truncated gaussian ϵ, u are drawn to satisfy the identifying rectangle restrictions using the probability integral transform method defined in Hajivassiliou (2008).

Hence, the model under conditions (i) and (ii) is coherent, since a complete DGP could be defined sequentially from the errors to the observables: $y_{i1}, w_{i1} \rightarrow y_{i2}, w_{i2} \rightarrow \cdots y_{i,t-1}, w_{i,t1} \rightarrow y_{it}, w_{it}$.

The proof for $\gamma \cdot \delta < 0$ because $\gamma > 0, \delta < 0$ is exactly symmetric and will not be repeated. ■

Appendix C: Monte-Carlo Evidence on CMLE

As we showed in Section 5 we obtain a coherent non-recursive model with interaction dummies included on both sides, provided we believe the feedback terms have opposite signs on the two sides. Note that it is sufficient to consider only the $\gamma \geq 0, \delta \leq 0$ case, since the reverse can always be subsumed by redefining both dependent binary variables to their complements $y'_{it} \equiv (1 - y_{it})$.

We performed extensive Monte-Carlo experiments designed to illustrate the consequences of adopting existing and novel estimation strategies for the problem of this paper. The experiments confirm that the CMLE approach under sign restrictions derived above provides reliable, consistent and efficient estimates of the underlying parameters including the two interaction terms. In contrast, the existing traditional approaches (unrestricted mle ignoring possible incoherency and mle that incorrectly assumes recursivity of the system) give seriously misleading and inconsistent results. The interested reader is referred to the online companion paper Hajivassiliou (2008) for an extensive presentation of the Monte-Carlos and detailed analysis and findings.[26] We give a brief summary here[27]:

- The Truncated CMLE proposed in this paper performs very satisfactorily, being the only consistent estimator for the reverse feedback cases, and with only small sacrifices in terms of efficiency in the recursive DGPs when it is not strictly necessary.
- Truncated CMLE also works well for the overlap region incoherency cases.
- Unrestricted likelihood estimation ignoring the resulting incoherency due to the empty or overlap region(s) is by far the worst performing estimator, dominated even by equation by equation univariate estimators which estimate the two equations separately while ignoring the other side of the model.

We plan to expand the online companion Monte-Carlo study Hajivassiliou (2008) to examine the relative performance of the two methods we developed here for addressing overlap region incoherency/incompleteness, namely the "contiguous triangles" and "truncated distributions" solutions.

[26] The cited study considered *nine* estimation approaches: 1. incorrectly forcing the old coherency condition to hold, i.e., assuming recursivity when in fact both feedback terms are present (estimators `e-trwn`=assuming $\delta = 0$ and `e-trnw` =assuming $\gamma = 0$); 2. unrestricted likelihood estimation, which ignores the resulting incoherency due to the empty or overlap region(s) (estimator `e-inco`); 3. restricted likelihood estimation conditioning on the data lying outside the empty region(s) of incoherency (estimators `e-sqpm` =assuming $(\gamma \geq 0, \delta \leq 0)$ and `e-sqmp`=assuming $(\gamma \leq 0, \delta \geq 0)$); 4. restricted likelihood estimation conditioning on the data lying outside the overlap region(s) of incoherency (estimators `e-sqpm` =assuming $(\gamma \geq 0, \delta \geq 0)$ and `e-sqmp`=assuming $(\gamma \leq 0, \delta \leq 0)$); and 5. lpols: (linear probability) ordinary least squares estimation of each binary probit equation ignoring the possible endogeneity of the interaction terms; and lp2sls: applying two-stage least squares recognizing that the two interaction terms on the rhs of each probit equation can be endogenous. In the cited study, six "true" models were generated, depending on whether interaction terms were allowed on one or both sides in each case, the nine estimators `e-trwn`, `e-trnw`, `e-inco`, `e-sqpm`, `e-sqmp`, `e-sqpp`, `e-sqmm`, `lpols`, and`lp2sls`were calculated.

[27] To simulate data from these six models studied in the Monte-Carlo paper, it was necessary to devise a methodology for generating standard gaussian variates truncated to lie outside an interval $\left[\underline{\lambda}, \bar{\lambda}\right]$. The following algorithm achieved this: let $z \sim n(0,1)$ and define $\tau \sim z \Big| \left\{ z \notin \left[\underline{\lambda}, \bar{\lambda}\right] \right\}$ then $cdf(z) : f(z) = \phi(z)$ and

$$cdf(\tau) : f(\tau) = \begin{cases} \dfrac{\phi(z)}{1 - \phi(\bar{\lambda}) + \phi(\underline{\lambda})} & \text{if } z < \underline{\lambda}, \\[2ex] \dfrac{\phi(\underline{\lambda})}{1 - \phi(\bar{\lambda}) + \phi(\underline{\lambda})} & \text{if } \underline{\lambda} < z \leq \bar{\lambda}, \\[2ex] \dfrac{\phi(z) - \phi(\bar{\lambda}) + \phi(\underline{\lambda})}{1 - \phi(\bar{\lambda}) + \phi(\underline{\lambda})} & \text{if } z > \bar{\lambda}. \end{cases}$$

The procedure is exact for a univariate z truncated on $\left\{ z \notin \left[\underline{\lambda}, \bar{\lambda}\right] \right\}$, but it will not work for higher dimensions. For DGPs with higher dimensions, the leading alternative procedures are Acceptance-Rejection and Gibbs Resampling — see Hajivassiliou and McFadden (1998) for discussion.

Appendix D: Detailed Estimation Results

Table 2
Innovation and Financing Constraints Joint Probit With Reverse Interaction Effects.

Full sample, nobs=1940		
	model 3	
	coeff.	std.
Innovation Equation		
constant	-7.235***	0.118
size	0.183***	0.020
market share	0.020	0.045
tp4	1.822***	0.183
tp3	1.0110***	0.199
tp2	0.437***	0.176
financial constraints	-0. 324**	0.255
11 industry dummies	misc	
Financial Constraint Equation		
constant	-1.221***	0.241
firm innovates	0.647***	0. 032
size	-0.016	0.073
collateral amount	0.030	0.050
banking debt ratio	0. 015***	0.002
own financing ratio	-0. 001***	0.001
profit margin	-0.002***	0.002
11 industry dummies	misc	
$corr_{12}$	-0.132***	0. 013
log lik innovation		
log lik fin constraint		
log lik bivariate	-1712	

Table 3
Innovation and Financing Constraints Joint Probit

With Reverse Interaction Effects and Dynamics Full sample, nobs=1512		
	Model 4	
	Coeff.	Std.
Innovation Equation		
Constant	-2.441***	0.323
Innov_{t-1}	0.829***	0.094
Size	0.256***	0.037
Market share	0.027	0.071
TP4	1.461***	0.201
TP3	0.932***	0.156
TP2	0.621***	0.143
Financial Constraints	-0. 447***	0.106
$\text{Financial Constraints}_{t-1}$	0. 300	0.123
11 Industry dummies	misc	
Financial Constraint Equation		
Constant	-0.885***	0.311
Firm Innovates_t	0.627***	0. 022
$\text{Firm Innovates}_{t-1}$	0. 236**	0.133
$\text{Financial constraints}_{t-1}$	0. 135***	0.093
Size	0.035	0.039
Collateral amount	0.003	0.002
Banking debt ratio	0.005	0.003
Own financing ratio	-0. 008***	0.002
Profit margin	-0.007***	0.002
11 industry dummies	misc	
$corr_{12}$	0.500**	0. 210
Log lik Innovation		
Log lik Fin Constraint		
Log lik Bivariate	-1331	

References

Aghion, P., Bloom, N., Blundell, R., Griffith, R., Howitt, P., 2005. Competition and innovation: An inverted-u relationship. The Quarterly Journal of Economics 120 (2), 701–728. https://doi.org/10.1093/qje/120.2.701.
Amemiya, T., 1973. Regression analysis when the dependent variable is truncated normal. Econometrica 41 (6), 997–1016.
Audretsch, D., 1995. Innovation, growth and survival. Int. J. Ind. Org. 13(4), 441–457.
Barghava, A., Sargan, J., 1982. Estimating dynamic random effects models from panel data covering short time periods. Econometrica 51, 1635–1660.
Beresteanu, A., Molchanov, I., Molinari, F., 2011. Sharp identification regions in models with convex moment predictions. Econometrica 79(6), 1785–1821.
Berry, S., Tamer, E., 2006. Microeconometric models of investment and employment. In: Blundell, R., Newey, T.P.W. (Eds.), Advances in Economics and Econometrics, Ninth World Congress of the Econometric Society. Cambridge, UK: Cambridge Univ. Press, pp. 46–85.
Blundell, R., Smith, R., 1994. Coherency and estimation in simultaneous models with censored or qualitative dependent variables. J. Econom. 64, 355–373.
Brown, J., Fazzari, S., Petersen, B., 2009. Financing innovation and growth: Cash flow, external equity, and the 1990s r&d boom. J. Finance 64, 151–185.
Brown, J., Martinsson, G., Petersen, B., 2012. Do financing constraints matter for r&d? Eur. Econ. Rev. 56(8), 1512–1529.
Chesher, A., Rosen, A., 2014. Simultaneous equations models for discrete outcomes: coherence, completeness, and identification. CeMMAP-UCL Working Paper.
Ciliberto, F., Tamer, E., 2009. Market structure and multiple equilibria in airline markets. Econometrica 77(6), 1791–1828.
Cohen, W.M., Klepper, S., 1996. Firm Size and the nature of innovation within industries: The case of process and product R&D. The Review of Economics and Statistics 78 (2), 232–243. https://doi.org/10.2307/2109925.
Cohen, W.M., Levinthal, D.A., 1989. Innovation and learning: The two faces of R&D. The Economic Journal 99 (397), 569–596. https://doi.org/10.2307/2233763.
Crépon, B., Duguet, E., Mairesse, J., 1998. Research, innovation and productivity: An econometric analysis at the firm level. Economics of Innovation and New Technology 7 (2), 115–158. https://doi.org/10.1080/10438599800000031.
Dagenais, M., 1997. A simultaneous probit model. GREQAM working paper.
de Paula, A., Tang, X., 2012. Inference of signs of interaction effects in simultaneous games with incomplete information. Econometrica 80 (1), 143–172.
Dubin, J., McFadden, D., 1984. An econometric analysis of residential electric appliance holdings and consumption. Econometrica 52(2), 345–362.
Falcetti, E., Tudela, M., 2008. What do twins share? A joint probit estimation of banking and currency crises. Economica 75, 199–221.
Gourieroux, C., Laffont, J., Monfort, A., 1980. Coherency conditions in simultaneous linear equations models with endogenous switching regime. Econometrica 48(3), 75–96.
Guiso, L., 1998. High tech firms and credit rationing. J. Econ. Behav. Org. 35, 39–59.
Hajivassiliou, V., Ioannides, Y., 2007. Unemployment and liquidity constraints. J. Appl. Econ. 22(3), 479–510.
Hajivassiliou, V., McFadden, D., 1998. The method of simulated scores for the estimation of ldv models. Econometrica 66(4), 863–896.
Hajivassiliou, V., 2007. Efficient and robust estimation of panel data ldv models with simultaneity, dynamics, and regressor-random effects correlation. LSE Department of Economics working paper.
Hajivassiliou, V., 2008. Novel approaches to coherency conditions in ldv models: Theory and monte-carlo experiments. LSE Department of Economics working paper, URL=http://econ.lse.ac.uk/staff/vassilis/papers/.
Hall, B., Lerner, J., 2010. The financing of r&d and innovation. Handbook of the Economics of Innovation. North-Holland, Boston, pp. 609–639.
Hanemann, W.M., 1984. Discrete/continuous models of consumer demand. Econometrica 52(3), 541–561.
Heckman, J., 1978. Dummy endogenous variables in simultaneous equation systems. Econometrica 66, 931–960.
Heckman, J., 1981. Dynamic discrete models. Structural Analysis of Discrete Data with Econometric Applications, pp. 179–195.Cambridge, Massachusetts
Heckman, J., 1981. Statistical models for discrete panel data. Structural Analysis of Discrete Data with Econometric Applications, pp. 114–178.Cambridge, Massachusetts
Heckman, J., Honoré, B., 1990. The empirical content of the roy model. Econometrica 58(5).
Holmstrom, B., 1989. Agency costs and innovation. J. Econ. Behav. Org. 12(3), 305–3271989.
Hottenrott, H., Peters, B., 2012. Innovative capability and financing constraints for innovation: More money, more innovation? Rev. Econ. Stat. 94(4), 1126–1142.
Klette, T., Kortum, S., 2004. Innovating firms and aggregate innovation. J. Polit. Econ. 112(5), 986–1018.
Lewbel, A., 2007. Coherency and completeness of structural models containing a dummy endogenous variable. Int. Econo. Rev. 48(4), November, 1379–1392.
Maddala, G., 1983. Limited-Dependent and Qualitative Variables inonometrics. Cambridge University Press.
Minetti, R., Murro, P., Rotondi, Z., Zhu, S., 2019. Financial constraints, firms' supply chains, and internationalization. J. Eur. Econ. Assoc. Volume 17, Issue 2, 327–375.
Mohnen, P., 2019. R&d, innovation and productivity. The Palgrave Handbook of Economic Performance Analysis.Belgrave, McMillan
Mohnen, P., Roller, L.-H., 2005. Complementarities in innovation policy. Eur. Econ. Rev. 49, 1431–1450.
Raymond, W., Mohnen, P., Palm, F., 2010. Sybrand Schim van der Loeff; Persistence of Innovation in Dutch Manufacturing: Is It Spurious? The Review of Economics and Statistics 92 (3), 495–504. https://doi.org/10.1162/REST_a_00004.
Savignac, F., 2008. The impact of financial constraints on innovation: What can be learned from a direct measure? Econ. Innovat. New Technol. 17(6), 553–569.
Tamer, E., 2003. Incomplete simultaneous discrete response model with multiple equilibria. Rev. Econ. Stud. 70, 147–165.
Uhlig, H., 2005. What are the effects of monetary policy on output? results from an agnostic identification procedure. J. Monetary Econ. 52 (2), 381–419.